7

SUBBAND IMAGE CODING

Bernd Girod, Frank Hartung, and Uwe Horn

Lehrstuhl für Nachrichtentechnik, Universität Erlangen-Nürnberg,
Cauerstraße 7, 91058 Erlangen, Germany

Visual information plays an important role in almost all areas of our life. Due to the vast amount of data associated with images, compression is a key technology for their digital transmission and storage. As an example, consider a single still color image of 512×512 pixels with 3 color components and 8 bits per sample per component resolution, which represents a data volume of 768 kBytes. For video, the problem is even more severe; a one-hour movie in a resolution according to the CCIR 601 studio standard [1] would need as much as 74.7 GBytes, when stored in uncompressed PCM format. These examples clearly demonstrate the need for efficient compression algorithms for images.

All methods for compression of still images or video are based on two fundamental principles. One principle is to exploit the properties of the signal source and to remove redundancy from the signal. The other principle is to exploit the properties of the signal receiver (usually the human visual system) and to omit parts or details of the signal that will not be noticed by the receiver. These principles are often referred to as redundancy reduction and irrelevancy reduction, respectively. The theory of subband decomposition provides an efficient framework for the implementation of schemes for redundancy and irrelevancy reduction.

It has been demonstrated repeatedly that subband and wavelet based schemes outperform other waveform based coding schemes [2] [3] [4]. Nevertheless, most of today's image and video coding standards use the discrete cosine transform (DCT) [5] for signal decomposition [6] [7] [8]. It can, however, be shown that the DCT, like all orthogonal transforms, is just a special case of subband decomposition [9]. Though the image coding standards are restricted to that special case, there is room for subband coding schemes in proprietary applica-

tions, which will become more important with the feasibility of software-only image codecs.

An important motivation for the use of subband decomposition schemes (rather than e.g. DCT-based schemes) is the demand for "scalable" image representations [10] [11] [12]. In the context of image coding, scalability means that the transmitted bit-stream can be decoded hierarchically, i.e., a low-resolution version of the transmitted image can be decoded with few operations, and the full resolution image will only be decoded if necessary or desired. This is for example useful for database browsing applications. For broadcasting applications, the low resolution signal can be transmitted with better error protection. This preserves at least a low resolution version of the image or video, if the channel characteristics vary, and the threshold behavior of digital transmission (where the receiver can decode the signal either in full digital quality or, if the bit error rate exceeds the threshold for the given error protection, nothing at all) is avoided [13]. Subband coding schemes are especially suitable for applications where scalability and "graceful degradation" are important issues.

This chapter treats the compression of still images by means of subband filter banks, whereas the next chapter is dedicated to subband compression of video. Section 1 reviews the theoretical foundations of subband coding. Rate-distortion theory is introduced, and rate-distortion functions for different signal models are presented. From rate-distortion theory, we can conclude that for stationary image models it is appropriate to decompose a signal into frequency bands and to encode those bands separately.

Section 2 discusses various approaches to subband decomposition of images. Equal band splitting and non-uniform band splitting are explained. The discrete cosine transform (DCT) is shown to be a special case of subband decomposition. Section 3 treats the compression of the subband signals after decomposition. Compression can be achieved by applying scalar quantization or vector quantization (VQ) to the subband coefficients, in combination with fixed or variable wordlength encoding. The concepts of scalar quantization, fixed-wordlength VQ, entropy-constrained VQ and lattice VQ are explained and compared experimentally. Optimal bit-rate allocation among the subbands is explained in subsection 3.4.

1 THEORETICAL FOUNDATIONS OF SUBBAND IMAGE CODING

Before we concern ourselves with specific subband image coding algorithms we would like to know how far we can expect to lower the transmission bit-rate for a given required picture quality. Are there fundamental limits that cannot be exceeded by any coding scheme? Such limits would be extremely helpful to judge the relative performance of a practical image coding scheme such as subband coding.

1.1 The Rate Distortion Function

Rate distortion theory is a branch of information theory that allows us to calculate performance bounds without consideration of a specific coding method. In particular, rate-distortion theory yields the minimum transmission bit-rate R, if a distortion D between the original image x at the transmitter and the reconstructed image $\hat{x}$ at the receiver shall not exceed a maximum acceptable distortion D^*. Unfortunately, the theory does not provide us with a method for constructing a practical optimum coder and decoder. We will see that rate-distortion theory can nevertheless provide us with very important hints about the properties of an optimum codec.

Rate distortion theory is based on two central concepts: "mutual information" and "distortion". Mutual information is a symmetric measure of the information that symbols x and $\hat{x}$ convey about each other. While the framework is entirely general, let's think of x representing an original image at the transmitter and $\hat{x}$ a reconstructed image at the receiver. We are mainly interested in average quantities in information theory. The average mutual information between ensembles X and $\hat{X}$, representing amplitude-continuous vector-valued random variables x and $\hat{x}$ with joint probability density function $p_{X\hat{X}}(x,\hat{x})$ and marginal probability density functions $p_X(x)$ and $p_{\hat{X}}(\hat{x})$ is defined as

$$I(X;\hat{X}) = \int_{\hat{x}}\int_{x} p_{X\hat{X}}(x,\hat{x}) \log \frac{p_{X\hat{X}}(x,\hat{x})}{p_X(x)p_{\hat{X}}(\hat{x})} dx\, d\hat{x}. \tag{7.1}$$

Mutual information is related to the differential entropy of an ensemble,

$$h(X) = \int_x p_X(x) \log \frac{1}{p_X(x)}\, dx, \tag{7.2}$$

and the conditional differential entropy of X given $\hat{X}$,

$$h(X|\hat{X}) = \int_{\hat{x}} \int_x p_{X\hat{X}}(x, \hat{x}) \log \frac{1}{p_{X|\hat{X}}(x|\hat{x})}\, dx\, d\hat{x} \tag{7.3}$$

by

$$I(X; \hat{X}) = h(\hat{X}) - h(\hat{X}|X) = h(X) - h(X|\hat{X}). \tag{7.4}$$

In (7.3), $p_{X|\hat{X}}(x|\hat{x})$ is the probability of an original image x, if a decoded image $\hat{x}$ is observed. Clearly, average mutual information is symmetric and non-negative

$$0 \leq I(X; \hat{X}) = I(\hat{X}; X). \tag{7.5}$$

The average mutual information $I(X; \hat{X})$ between the original image at the coder and the reconstructed image at the decoder is related to the channel capacity C available between X and $\hat{X}$. The channel capacity is the maximum number of bits per symbol for a given symbol-rate that a transmission channel can accommodate without bit-errors. The channel capacity can be shown to be the maximum of the average mutual information between transmitter and receiver [14], i.e.

$$I(X; \hat{X}) \leq C. \tag{7.6}$$

The distortion measure $d(x, \hat{x})$ is a scalar quantity that should reflect the fidelity of the reproduction of an original image x by an image $\hat{x}$. To qualify as a proper distortion measure, we require

$$d(x, \hat{x}) \geq 0 \tag{7.7}$$

with equality if $x = \hat{x}$. The average distortion then is

$$D = \int_{\hat{x}} \int_{x} p_{X\hat{X}}(x, \hat{x}) d(x, \hat{x}) \, dx \, d\hat{x}. \tag{7.8}$$

The 'rate-distortion function' is defined as

$$R(D^*) = \inf_{p_{\hat{X}|X}} \{I(X; \hat{X}) : D \leq D^*\}, \tag{7.9}$$

i.e., $R(D^*)$ is the greatest lower bound of the average mutual information, subject to the constraint that the average distortion D may not exceed D^*. When there is no risk of confusion, we will drop the asterisk in the following and write $R(D)$ for the rate-distortion function. It can be shown that $R(D)$ is a monotonically decreasing convex function [14][15]. At $R = 0$, it assumes a maximum distortion $D_{\max}$. For an amplitude-continuous source, the rate required for an exact reproduction $\hat{x} = x$ is infinite, nevertheless the rate $R(0)$ can be finite, if the distortion measure contains a threshold below which $d(x, \hat{x}) = 0$, even if $x \neq \hat{x}$. Since $R(D)$ is monotonically decreasing, we can also use the distortion rate function $D(R)$ when it is more convenient.

The rate-distortion function is a performance bound that no source coder can beat. Conversely, rate-distortion theory shows that a source coder with a performance arbitrarily close to the rate-distortion function exists. The typical theoretically optimum source coder encodes a very large number of symbols jointly, thus requiring a very large memory and introducing a very large delay. While this might be impractical, it suggests that a good coder jointly encodes many symbols. We will make use of this idea in the following.

1.2 Shannon Lower Bound

The Shannon lower bound is a useful lower bound of the rate-distortion function. With a simple coordinate transformation in (7.3), it can be shown that

$$h(X - \hat{X}|\hat{X}) = h(X|\hat{X}), \tag{7.10}$$

where $X - \hat{X}$ is the reconstruction error ensemble representing the amplitude-continuous vector-valued difference $x - \hat{x}$ [15]. With this, we can rewrite the rate-distortion function as

$$\begin{aligned} R(D^*) &= \inf_{p_{\hat{X}|X}} \{h(X) - h(X|\hat{X}) : D \leq D^*\} \\ &= h(X) - \sup_{p_{\hat{X}|X}} \{h(X|\hat{X}) : D \leq D^*\} \end{aligned}$$

$$= \quad h(X) - \sup_{p_{\hat{X}|X}} \{h(X - \hat{X}|\hat{X}) : D \leq D^*\}. \tag{7.11}$$

Observing that

$$h(X - \hat{X}|\hat{X}) \leq h(X - \hat{X}) \tag{7.12}$$

we arrive at the Shannon lower bound

$$R(D^*) \geq h(X) - \sup_{p_{\hat{X}|X}} \{h(X - \hat{X}) : D \leq D^*\}. \tag{7.13}$$

Equality in (7.12) and (7.13) holds for statistical independence between $X - \hat{X}$ and $\hat{X}$. Thus, ideally, the source coding scheme would introduce a reproduction error $x - \hat{x}$ that is statistically independent of the reconstructed signal $\hat{x}$. Note that this is not always possible, particularly not at low rates. Nevertheless, it provides another guideline for the design of an efficient coding scheme.

It is possible to draw additional conclusions from the Shannon lower bound, if the distortion measure is given. Let us consider a single letter distortion measure

$$d = (x - \hat{x})^2, \tag{7.14}$$

i.e., the squared difference between the original and reconstructed images is calculated on a sample-by-sample basis. With the mean squared error $D \leq D^*$, the differential entropy of the reconstruction error is bounded according to [16]

$$h(X - \hat{X}) \leq \frac{1}{2} \log(2\pi e D^*). \tag{7.15}$$

Equality in (7.15) holds for a Gaussian pdf of $x - \hat{x}$. Successive values of $x - \hat{x}$ should all be independent, identically distributed random variables. Thus, an optimum source coder for a mean squared error distortion measure should produce white Gaussian noise independent from the reconstructed signal.

1.3 Memoryless Gaussian Source

In general, it is difficult to calculate the rate-distortion function. However, there are some important cases where the result can be stated analytically. Consider a memoryless Gaussian source with variance σ^2. For this case, the rate-distortion function is

$$R(D) = \frac{1}{2} \max\{\log \frac{\sigma^2}{D}, 0\}. \tag{7.16}$$

If $\log = \log_2$ in (7.16), the unit for the rate is "bits". The rate-distortion curve is plotted as curve "ML" in Fig. 1, with distortion shown on a logarithmic axis. With signal-to-noise ratio defined as

$$\mathrm{SNR} = 10 \log_{10} \frac{\sigma^2}{D} \,\mathrm{dB} \tag{7.17}$$

we encounter the information theoretic rule-of-thumb that 6 dB corresponds to 1 bit. Remarkably, the rate-distortion curves for non-Gaussian sources with the same variance σ^2 are always below the Gaussian $R(D)$ curve. Thus, Gaussian sources are the most demanding sources for coding.

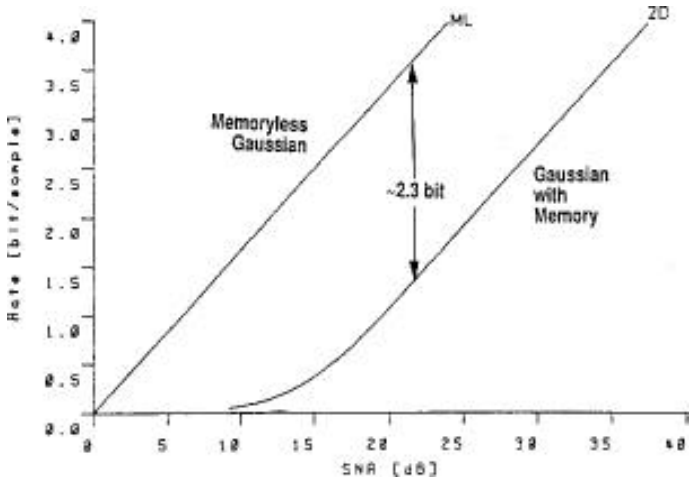

Figure 1 Rate distortion function for a memoryless Gaussian source (ML) and for a Gaussian source with power spectral density as shown in Fig. 3 (2D).

1.4 Gaussian Source with Memory

For sources with memory, the correlation between neighboring samples can be exploited and a lower rate can be achieved. Let us consider an ergodic two-dimensional space and amplitude continuous Gaussian source with power spectral density $\Phi_{xx}(\Omega_1, \Omega_2)$. Again, we use the squared error distortion measure (7.14). While the rate-distortion function cannot be given in closed form for this source and distortion measure, it can be stated in parametric form as

$$D(\theta) \quad = \quad \frac{1}{4\pi^2} \int_{\Omega_2} \int_{\Omega_1} \min(\theta, \Phi_{xx}(\Omega_1, \Omega_2)) d\Omega_1 \, d\Omega_2 \tag{7.18a}$$

$$R(\theta) \quad = \quad \frac{1}{8\pi^2} \int_{\Omega_2} \int_{\Omega_1} \max(0, \log \frac{\Phi_{xx}(\Omega_1, \Omega_2)}{\theta}) d\Omega_1 \, d\Omega_2. \qquad (7.18b)$$

Each value of θ in (7.18) (within an appropriate range) produces a point $R(D)$ of the rate-distortion curve. Again, for non-Gaussian sources with the same power spectral density, the rate-distortion curve is always below the Gaussian case.

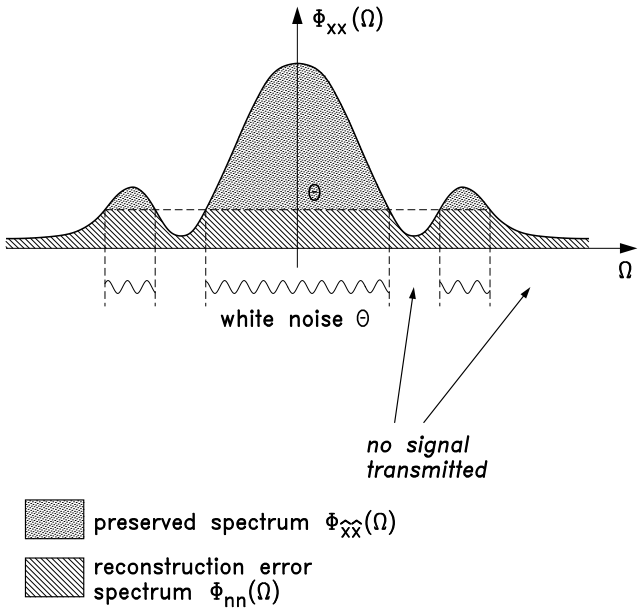

Figure 2 Interpretation of the rate-distortion function for a one-dimensional Gaussian source with memory.

The parametric solution (7.18) can be interpreted as illustrated in Fig. 2 for a one-dimensional case. Let $\Phi_{nn}(\Omega_1, \Omega_2)$ denote the power spectral density of the reconstruction error $x - \hat{x}$ (compare Fig. 23). In the frequency range where $\Phi_{xx}(\Omega_1, \Omega_2) > \theta$, white noise with power spectral density $\Phi_{nn}(\Omega_1, \Omega_2) = \theta$ is introduced. With the same arguments as used in the context of the Shannon lower bound, it can be shown that this noise is uncorrelated with the reconstructed signal $\hat{x}$. Thus

$$\Phi_{\hat{x}\hat{x}}(\Omega_1, \Omega_2) = \Phi_{xx}(\Omega_1, \Omega_2) - \theta, \forall \Omega_1, \Omega_2 : \Phi_{xx}(\Omega_1, \Omega_2) > \theta. \qquad (7.19)$$

In the frequency range where $\Phi_{xx}(\Omega_1, \Omega_2) < \theta$ the signal power itself is smaller than the coding noise introduced. Thus, it leads to the minimum mean squared error to transmit no signal at all in these parts of the spectrum, and

$$\left.\begin{array}{rcl} \Phi_{\hat{x}\hat{x}}(\Omega_1, \Omega_2) & = & 0 \\ \Phi_{nn}(\Omega_1, \Omega_2) & = & \Phi_{xx}(\Omega_1, \Omega_2) \end{array}\right\} \forall \Omega_1, \Omega_2 : \Phi_{xx}(\Omega_1, \Omega_2) < \theta \tag{7.20}$$

Since the overall rate in (7.18b) is the integral over the rate contributions dR of individual frequency components, an optimum coder can be built by splitting the original signal into spectral components of infinitesimal bandwidth $d\Omega_1\, d\Omega_2$ and encoding these spectral components independently [17]. This suggests an approach where the subband components with energy above a threshold are encoded with a number of bits that is proportional to the logarithm of their energy, while the remaining subbands are suppressed.

1.5 R(D) for a Gaussian Image Model with Exponentially Decaying Autocorrelation Function

We can compute the rate-distortion function (7.18), if we know the power spectrum $\Phi_{xx}(\Omega_1, \Omega_2)$. A commonly used model for the autocorrelation function of an image is the isotropic, exponentially decaying autocorrelation function

$$R_{xx}(\Delta_1, \Delta_2) = e^{-\Omega_0 \sqrt{{\Delta_1}^2 + {\Delta_2}^2}}, \tag{7.21}$$

where Δ_1 and Δ_2 are the horizontal and vertical distances between samples of x, and Ω_0 is a constant. The corresponding power spectral density is computed as the 2D Fourier transform of (7.21),

$$\Phi_{xx}(\Omega_1, \Omega_2) = \frac{2\pi}{\Omega_0^2}\left(1 + \frac{\Omega_1^2 + \Omega_2^2}{\Omega_0^2}\right)^{-\frac{3}{2}}. \tag{7.22}$$

The model power spectrum (7.22) is shown in Fig. 3 with parameters adapted to a video signal sampled in the "Common Intermediate Format" (CIF, 288 lines $\times$ 352 pixels per line for a video frame with 4/3 aspect ratio). The rate-distortion curve is computed by numerically integrating (7.18). The result is shown in Fig. 1 as curve "2D". The Gaussian source with a power spectrum (7.22) requires about 2.3 bits/sample less than the memoryless Gaussian source (7.16) at high rates. The slope is again 6 dB/bit at high rates where the rate-distortion curve coincides with its Shannon lower bound.

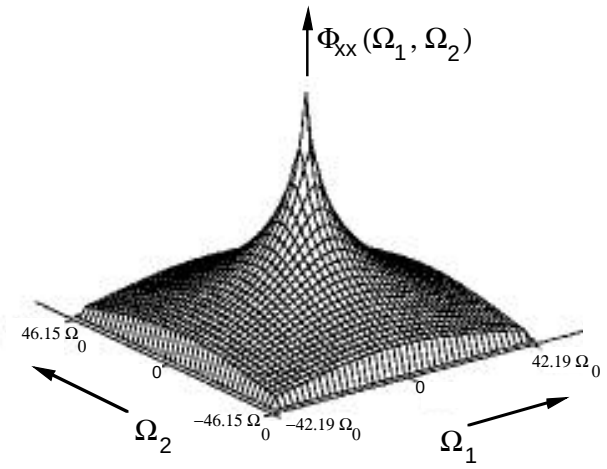

Figure 3 Power spectral density model of an image (vertical axis in logarithmic scale).

2 IMAGE SUBBAND DECOMPOSITION

Rate distortion theory suggests that an efficient source coder splits the image into frequency bands and independently encodes the individual subband signals. While images are two-dimensional signals, we discuss one-dimensional subband coding first. Most two-dimensional subband decompositions are built by cascading one-dimensional subband filter banks.

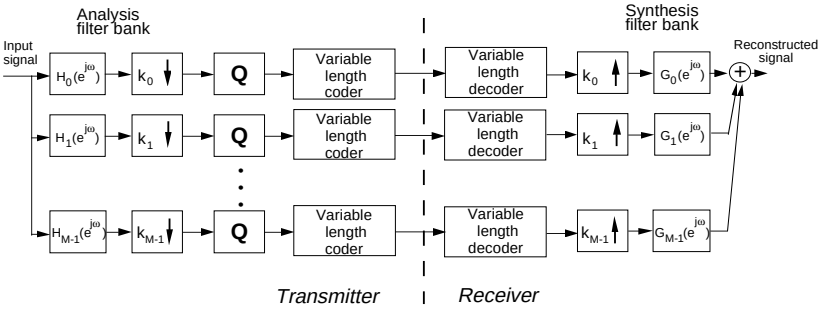

Figure 4 Diagram of a one-dimensional M-channel subband coding scheme.

A general one-dimensional subband coding scheme is shown in Fig. 4. The sampled input signal is convolved in a bank of M bandpass analysis filters with frequency responses $H_i(e^{j\omega})$ and down-sampled by factors k_i that correspond to the bandwidths of $H_i(e^{j\omega})$. The frequency responses and down-sampling factors are usually chosen such that the number of samples before and after subband decomposition is the same, i.e.,

$$\frac{1}{k_0} + \frac{1}{k_1} + \ldots + \frac{1}{k_{M-1}} = 1. \tag{7.23}$$

Eq. (7.23) ensures that there is no additional redundancy introduced by the subband decomposition. The subband signals are quantized and transmitted using an appropriate fixed or variable length code. At the decoder, the subband signals are decoded, up-sampled by factors k_i and passed through a bank of synthesis filters with frequency responses $G_i(e^{j\omega})$. The output signals of the synthesis filters are finally summed up to yield the reconstructed signal.

2.1 Subband Filter Banks

As discussed in chapters 1–4, there are many filter banks available for general use. However, when choosing filters for subband image decomposition, there are additional requirements that are specific to image coding [18]. Analysis filters should have a short impulse response to preserve the localization of image features. Synthesis filters should also have a short impulse response in order to prevent spreading of artefacts resulting from quantization errors at edges and other local features. Long synthesis filters often have very good mean squared error performance but lead to annoying ringing effects around edges. In addition, linear phase filters are desirable for subband image coding. Filters with nonlinear phase introduce subjectively unpleasant waveform distortions, when the lowpass channel is viewed by itself.

Many filters have been suggested for subband image coding [19] [20] [21] [22] [23] [24] [25]. We have used the 9-tap QMFs

$$\begin{aligned} H_0(e^{j\omega}) = G_0(e^{j\omega}) &= 0.5645751 \\ &+ 0.2927051(e^{j\omega} + e^{-j\omega}) \\ &- 0.05224239(e^{j2\omega} + e^{-j2\omega}) \\ &- 0.04270508(e^{j3\omega} + e^{-j3\omega}) \\ &+ 0.01995484(e^{j4\omega} + e^{-j4\omega}). \end{aligned} \tag{7.24}$$

The highpass filters $H_1(e^{j\omega})$ and $G_1(e^{j\omega})$ are given by $H_0(e^{j(\omega+\pi)})$ and $-G_0(e^{j(\omega+\pi)})$ respectively.

The reconstruction error introduced by the QMF bank is at -50 dB and thus smaller than the typical quantization error. The stopband attenuation of the analysis and synthesis filters is about 35 dB.

It should be noted that these filters require that the subsampling points be staggered between the two subbands, i.e., the low band is sampled at positions $0, 2, 4, \ldots$, and the high band is sampled at the positions $1, 3, 5, \ldots$. The alternating sampling preserves the information more uniformly and gives better reconstruction than even length QMFs with coinciding sampling in both bands.

2.2 Equal Band Splitting

Subband decomposition of images into more than two bands can conveniently and efficiently be carried out by cascading horizontal and vertical two-band filter banks. As an example, Figs. 5 and 6b illustrate a decomposition into 4×4 subbands. Any two-band filter bank, such as the short-kernel filters from [23] or the QMFs (7.24) can be employed in this structure. Note that the filter bank in Fig. 5 is separable in horizontal and vertical directions, i.e. the resulting frequency responses of the individual subbands can be written as a product of 1D functions of ω_1 and ω_2.

Rather than cascading two-band filter banks, we can also use an M-channel filter bank for image subband decomposition. An example of a separable filter bank with $M \times M$ channels is the Lapped Orthogonal Transform [26]. The majority of filter banks investigated for image coding are separable into horizontal and vertical filter banks. Such filter banks perform better on strictly horizontal and vertical edge features than on diagonal ones. While horizontal and vertical edges fall into distinct subbands in decompositions such as the one shown in Figs. 5 and 6b, diagonal edges with $45°$ and $135°$ orientations fall into the same subbands. Non-separable subband filter banks with directional characteristics have been investigated by [27] [28]. However, the computational effort required for non-separable filter banks is often prohibitively large.

2.3 Discrete Cosine Transform as a Subband Coding Technique

The discrete cosine transform (DCT), introduced by Ahmed, Natarajan, and Rao in 1974 [5], [29] has been adopted in several image coding standards in

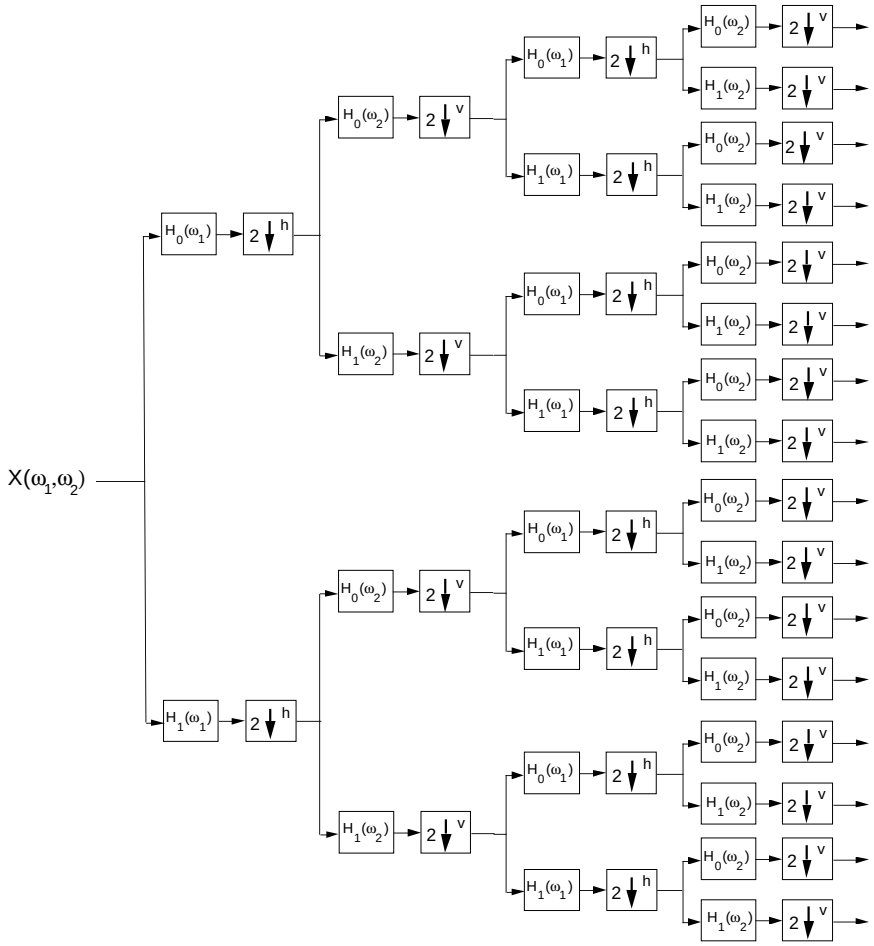

Figure 5 4×4-band analysis filter bank built out of two-band filter banks. $2{\downarrow}^h$ denotes horizontal subsampling by a factor of 2, $2{\downarrow}^v$ denotes vertical subsampling by a factor of 2.

the last years. We will show in this section that the DCT can be interpreted as a special subband coding technique. In transform image coding, the image is subdivided into square blocks, and each block undergoes an orthonormal transformation $\mathbf{A}$ (Fig. 7). We can write the transformation as a matrix multiply

$$\mathbf{c} = \mathbf{A}\mathbf{x} \tag{7.25}$$

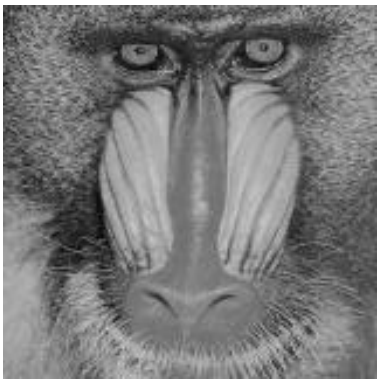

Figure 6a Original image Mandrill.

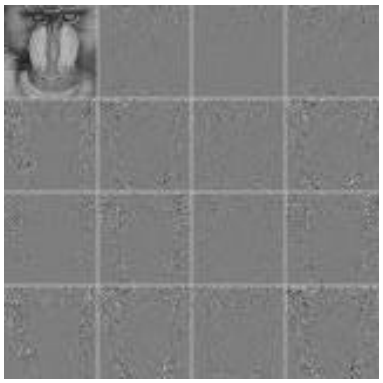

Figure 6b Mandrill decomposed into 4×4 bands (using short-kernel filters from [23]).

where $\mathbf{x}$ is a column vector formed by the $M \times M$ samples of an image block, $\mathbf{A}$ is a $M^2 \times M^2$ matrix, and $\mathbf{c}$ is the vector of M^2 transform coefficients. For an orthonormal transform,

$$\mathbf{x} = \mathbf{A^{-1}c} = \mathbf{A^T c} \tag{7.26}$$

i.e., the inverse transform matrix is the transpose of the forward transform matrix. The equation $\mathbf{x} = \mathbf{A^{-1}c}$ can be interpreted as a representation of an image block by a superposition of "basis functions" (columns of $\mathbf{A}^{-1}$) which are weighted by coefficients c_i (elements of vector $\mathbf{c}$).

A two-dimensional $M \times M$ DCT is separable into a horizontal and a vertical one-dimensional M-DCT. The elements a_{ik} of the transform matrix $\mathbf{A}$ for the DCT ("DCT-II" after the classification in [29]) of length M are

$$a_{ik} = \alpha_i \cos \frac{\pi(2k+1)i}{2M} \qquad \text{with} \qquad \alpha_0 = \sqrt{\frac{1}{M}}, \quad \alpha_i = \sqrt{\frac{2}{M}} \quad \forall i \neq 0. \tag{7.27}$$

Fig. 8 shows the 64 basis functions for an 8×8 DCT.

The DCT is closely related to the discrete Fourier transform (DFT). It can be shown that the DFT is nearly optimum for images with stationary signal statistics when the block size is very large [30]. The argument is similar to the rate-distortion analysis presented in section 1.4 that provided the insight that ergodic sources should be decomposed and coded in independent subbands. For

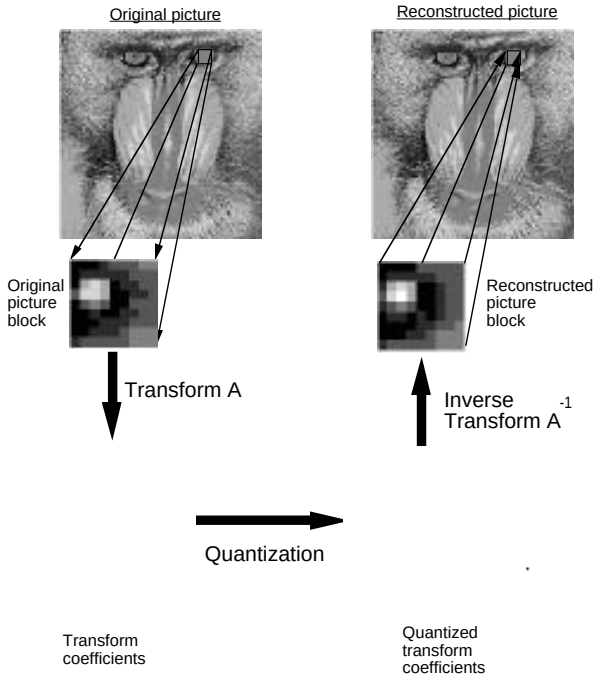

Figure 7 Principle of transform coding.

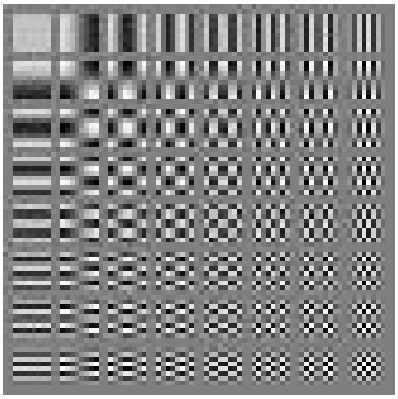

Figure 8 Basis functions of an 8×8 DCT.

finite block sizes, block boundary effects with DFT coding are very annoying owing to the circular topology underlying the DFT. Since, with a DFT, the right edge of a block is "connected" to the left edge and the top edge to the bottom edge, the signal discontinuity at the block edge leads to severe Gibbs ringing when higher DFT coefficients are suppressed in a coder. This problem is greatly reduced by mirroring an $M \times M$ block horizontally and vertically such that a $2M \times 2M$ block results (Fig. 9).

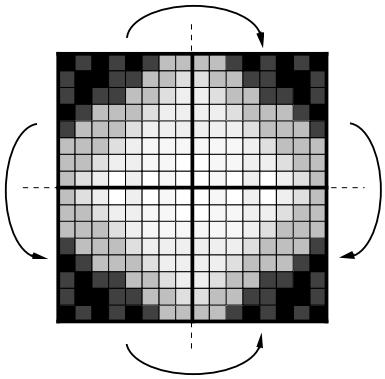

Figure 9 Horizontal and vertical block mirroring leads to an interpretation of the DCT in terms of the DFT.

The $2M \times 2M$ block is then transformed using the DFT. Due to the symmetry of the mirrored block, the transform has only $M \times M$ degrees of freedom and only $M \times M$ non-zero coefficients remain. They are the coefficients of a new transform, the discrete cosine transform, with the transform matrix $\mathbf{A}$ (7.27).

An $M \times M$ DCT can be viewed as a subband decomposition into $M \times M$ subbands. The impulse responses of the analysis and synthesis filters corresponding to an M-point DCT are exactly M samples long. In the analysis filter bank, the image is convolved separately with each row of the transform matrix $\mathbf{A}$ (7.27), and then each subband signal is downsampled by a factor $k_i = M$. Values that are omitted in the down-sampling stage need not be calculated in the analysis filter convolutions. Since the impulse responses are not longer than the down-sampling factors, the input signal can be subdivided into non-overlapping portions and processed in independent blocks. A similar argument applies to the inverse DCT as a synthesis filter bank. Note that the DCT is in fact a perfect reconstruction filter bank with identical analysis and synthesis filters, except for a horizontal and vertical reversal of the filter impulse responses. Fig. 10 shows the frequency response of a forward (and,

at the same time, an inverse) 8-point DCT when viewed as a filter bank. The frequency selectivity of the DCT is remarkably poor. On the other hand, the compact support of the filter impulse responses is a very desirable property. Exact image subband decomposition and reconstruction using the DCT are described in [31]. Approaches combining the DCT and subband filter banks are described in [32] [33].

2.4 Non-Uniform Band Splitting

When a signal is spectrally flat within a band, memoryless coding is optimal. Fig. 3 shows the model of a typical image power spectral density (7.22). Apparently, the spectrum varies more rapidly at low frequencies than at high frequencies. When decomposing the spectrum (7.22) into equal bands, the high bands will be spectrally flatter than the low bands. This suggests a non-uniform band splitting that decomposes the lower frequency part into narrower bands than the higher frequency range.

A signal decomposition into infinitesimally narrow subbands is suggested by the rate-distortion theoretical argument in Section 1.4. This argument applies to ergodic Gaussian sources. Images, however, contain spatially localized features, such as edges, and very narrow subbands do not yield good coding results. The impulse responses of filters producing narrow subbands are necessarily long, and the localization of image features is poor after filtering. In principle, the same features can occur in an image on all scales, due to the perspective projection of objects at all distances from the camera onto the image plane. This suggests using short filter impulse responses for high frequency channels and long impulse responses for low frequency channels.

A powerful non-uniform subband decomposition is the subband pyramid illustrated in Fig. 11. The image spectrum is split into four bands containing horizontal lows/vertical lows (LL), horizontal lows/vertical highs (LH), horizontal highs/vertical lows (HL), and horizontal highs/vertical highs (HH). This four-band decomposition can be accomplished by cascading a horizontal and a vertical two-band filter bank. In the next stage, the same four band decomposition is applied to the LL component only. The LH, HL, and HH components are left alone. The procedure is repeatedly applied to each resulting low frequency band, resulting in a decomposition into octave bands.

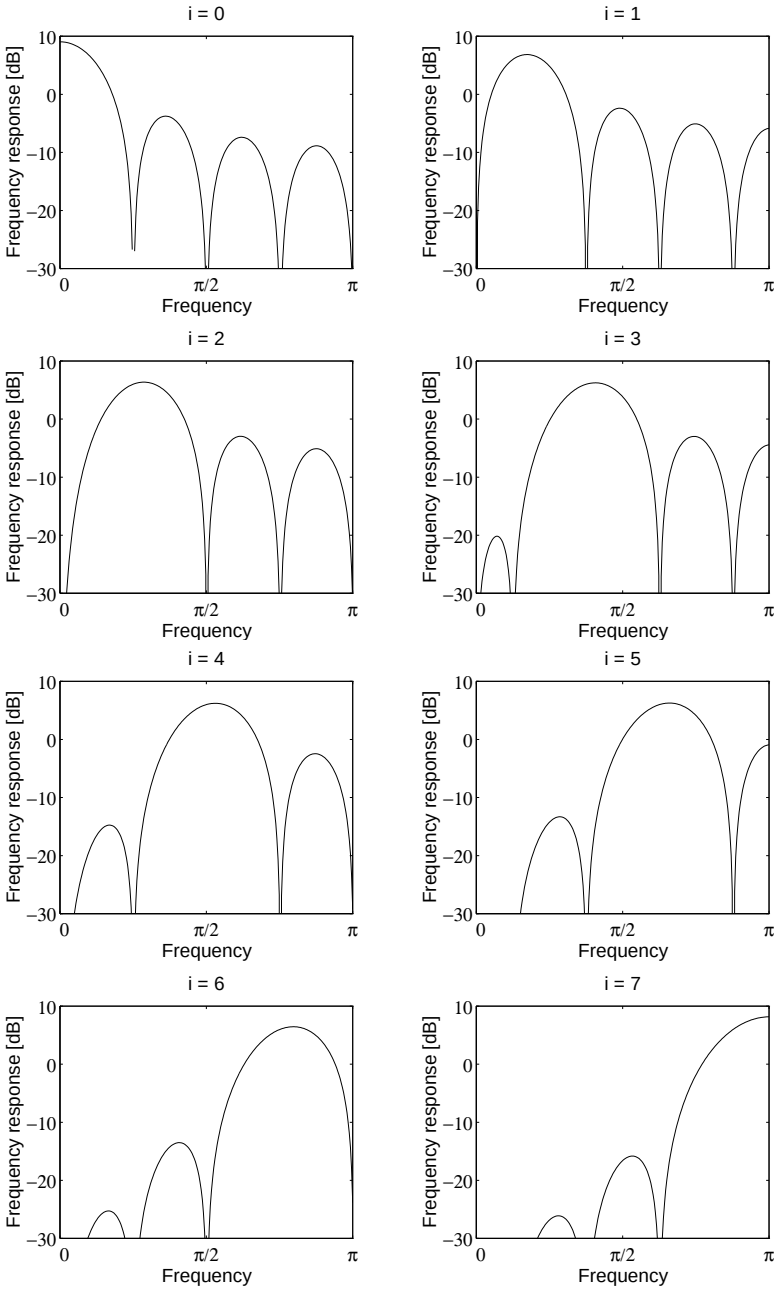

Figure 10 Subband frequency response of 8-point DCT.

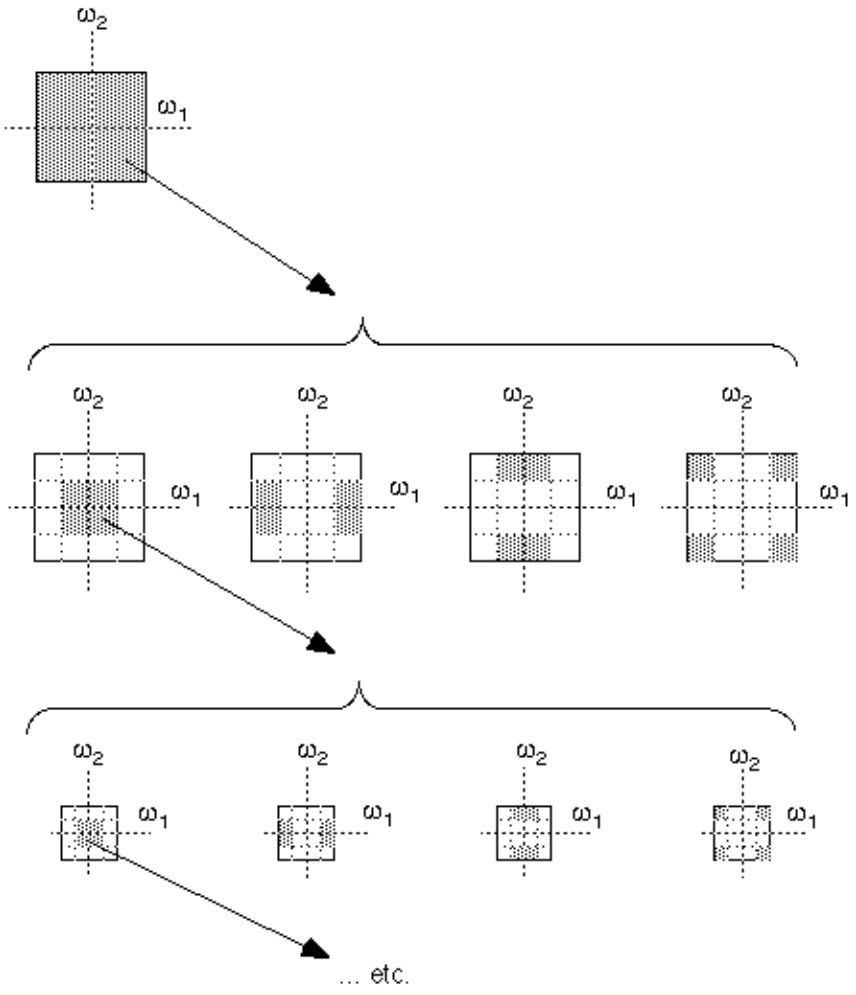

Figure 11 Subband pyramid decomposition in the frequency domain.

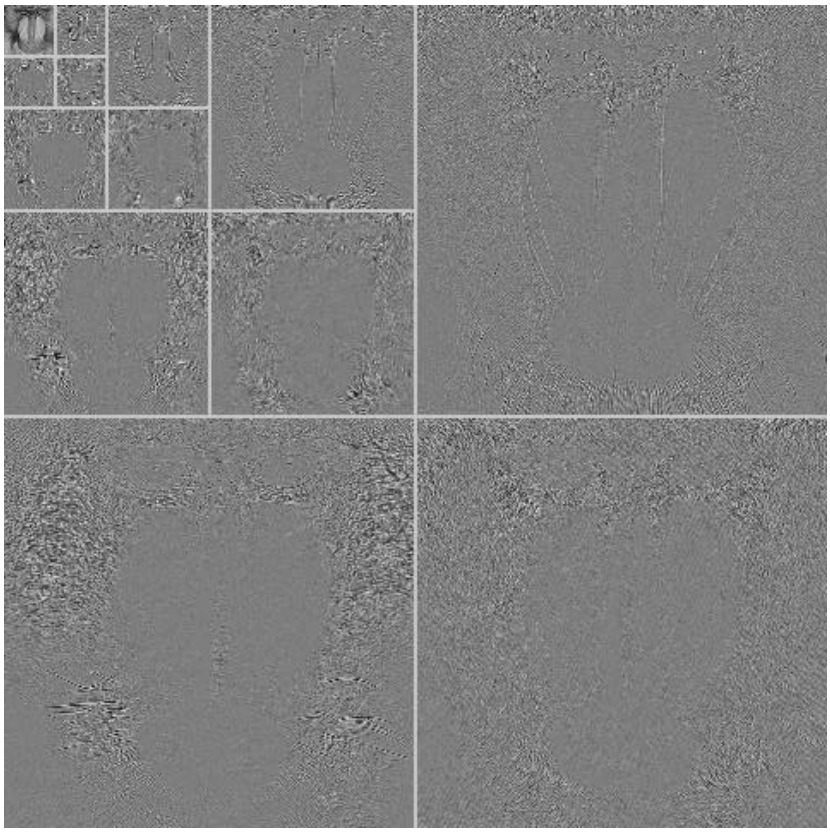

Figure 12 Image Mandrill (Fig. 6a) decomposed by four level subband pyramid.

Fig. 12 shows an image decomposed into a subband pyramid with four levels. The QMFs (7.24) were used for the two-band decomposition at each level of the subband pyramid.

Besides critically sampled subband pyramids (Eq. 7.23), oversampled pyramid decompositions can be employed for image coding. They were first introduced as bandpass pyramids by Burt and Adelson in 1983 [34] and successively refined, e.g. by [10] [13]. An input picture is first lowpass filtered and downsampled. By interpolating the low resolution image back to its original resolution, a prediction of the original image is obtained. The resulting prediction error image is a bandpass signal containing the missing detail of the prediction from the lower resolution. This decomposition into a lower resolution image and a prediction error image can be repeated for the lower resolution image to obtain more pyramid layers. Note that the number of samples increases by up to 1/3 for 2 : 1 two-dimensional subsampling. On the other hand, one gains complete freedom in choosing appropriate filters. Lower resolution images within predictive resolution pyramids often have better subjective image quality than those obtained from critically sampled subband pyramids.

3 COMPRESSION OF IMAGE SUBBAND SIGNALS

After the decomposition stage, the image is split into subimages that contain spectral components of the original image. If the subimages are critically sampled, the total number of samples is the same as in the original image. The decomposition is still fully reversible, and no compression has been accomplished. Quantization can be performed on single subband coefficients [35] [36] [37] ("scalar quantization"), or on several coefficients together ("vector quantization (VQ)"). Vector quantization techniques have been proved to be very powerful for quantization of subbands [32] [38] [39] [40] [41] [42] [43] [44]. In subsection 3.1, scalar quantization is explained. In subsection 3.2, different approaches to vector quantization are presented. In subsection 3.3, the performance of different quantization schemes is illustrated by coding experiments on images. Finally, subsection 3.4 discusses bit allocation among subbands.

3.1 Scalar Quantization

A scalar quantizer rounds each sample of the input x independently to the nearest representative output level. Fig. 13 shows an example of the staircase input-output characteristic of a scalar quantizer. The input signal amplitude within a certain range is mapped onto one common representative output level. The levels that separate the input signal ranges are called decision thresholds. Simultaneously finding the N representative levels and $N-1$ decision thresholds of a quantizer yielding minimum mean squared quantization error for a given probability density function (pdf) $p_X(x)$ of the input signal x is a classical problem that has been solved by Lloyd [45] and Max [46] independently. The solution, the "Lloyd–Max quantizer," has the following properties:

- The decision thresholds lie exactly half-way between the representative levels.
- The representative levels are the centroids of the pdf between successive decision thresholds.

Since the optimum decision thresholds depend on the representative levels, while the optimum representative levels depend on the decision thresholds, a closed form solution of the Lloyd–Max quantization problem can usually not be found except for some trivial cases. Iterative schemes have to be used that converge to the optimum solution.

For sufficiently fine quantization and a smooth pdf, an approximate solution of the Lloyd–Max quantization problem can be obtained according to

$$\Delta\hat{x}(x) = c \cdot (p_X(x))^{-\frac{1}{3}} \tag{7.28}$$

where $\Delta\hat{x}(x)$ is the distance between two successive quantizer representative levels. Interestingly, the approximation (7.28) was proposed by Panter and Dite [47] before the exact solution was known. The Panter and Dite approximation also provides an estimate of the resulting quantization error variance

$$\sigma_q^2 = \frac{1}{12N^2}\Big(\int_x (p_X(x))^{\frac{1}{3}} dx\Big)^3 \tag{7.29}$$

where N is again the number of representative levels.

The Lloyd–Max quantizer and the Panter and Dite approximation minimize the mean squared quantization error for a given number of representative levels.

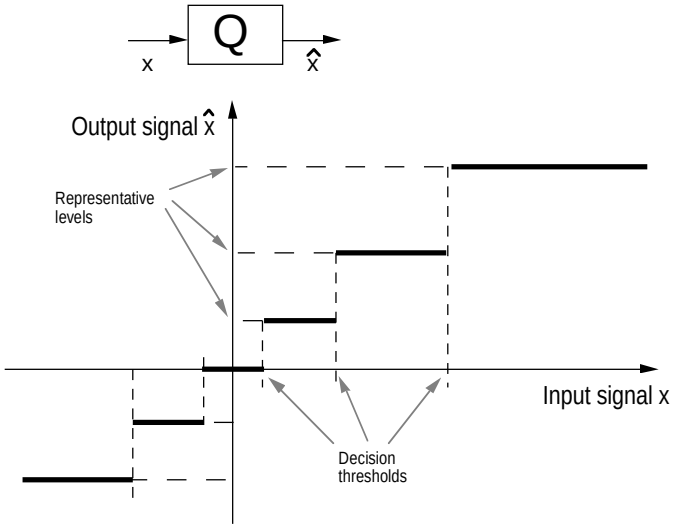

Figure 13 Example for a quantizer input-output characteristic.

The solution implies that each representative level is coded by a fixed codeword length. Often, we want to employ an entropy coder after the quantizer which generates codewords of variable length in order to approach the entropy of the quantized signal. In this case, it is not appropriate to minimize mean squared quantization error subject to a fixed number of representative levels, but rather subject to a fixed entropy of the quantized signal $\hat{x}$. It can be observed that a uniform quantizer often yields a better result than a Lloyd–Max quantizer, when followed by entropy coding. In fact, for sufficiently fine quantization, a uniform quantizer yields minimum entropy [48].

Scalar quantization followed by entropy coding is often used in the context of subband image coding [35] [36] [37]. The amplitude histogram of a typical subband signal is shown in Fig. 14. It can be approximated well by a Laplacian pdf. Optimum scalar quantizers for Laplacian pdfs and Gamma pdfs have been investigated extensively by Brusewitz [49]. His work shows that uniform quantizers with a representative level at zero and a larger threshold around zero yield excellent results even at low rates (Fig. 15). Uniform threshold quantizers are employed today in the DCT-based image coding standards ITU-T H.261, MPEG, and JPEG [6] [7] [8].

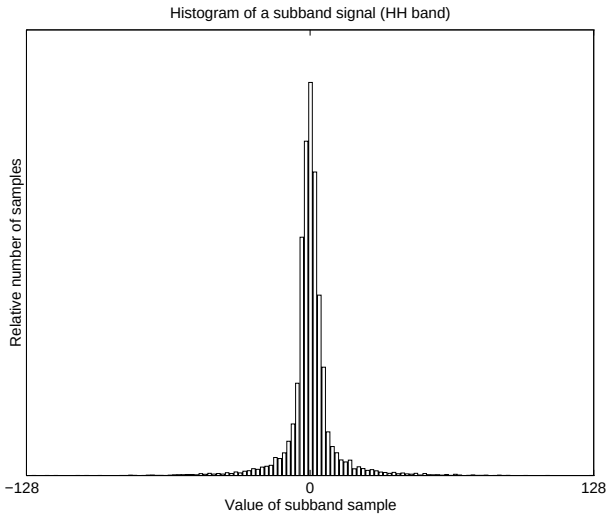

Figure 14 Histogram of a subband signal.

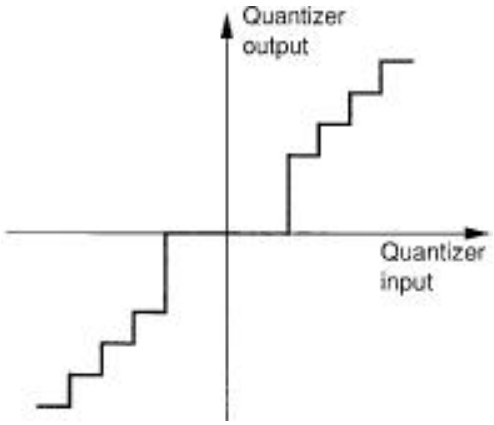

Figure 15 Uniform threshold quantizer.

3.2 Vector Quantization Techniques for Image Subband Coding

A subband decomposition reduces or even eliminates the correlation between image samples, but there are still statistical dependencies of higher order within or across subband signals unless the signal statistics are Gaussian. A vector quantizer, which jointly quantizes several samples, can also exploit these statistical dependencies of higher order. The input to a vector quantizer can be formed by combining several neighboring samples within an image subband [40], or spatially coinciding samples from different subbands [38], or a combination of both. In [42] several cases for forming vectors within a subband coding scheme for color images are investigated. In this section, we introduce some basic vector quantization (VQ) techniques as well as lattice vector quantization. A broader overview can be found in [50] [51].

Fixed-Wordlength VQ

In 1980, Linde, Buzo, and Gray generalized the Lloyd algorithm for the design of nonuniform scalar quantizers to yield a codebook for vector quantizers containing a fixed number of codevectors [52]. Their algorithm, known as the "LBG algorithm" or "generalized Lloyd algorithm", computes a codebook with minimum average distortion for a given training set and given codebook size. In the following, let us assume squared error as the distortion measure, which is, for k-dimensional vectors $\mathbf{x}$ and $\mathbf{y}$, defined as

$$d(\mathbf{x}, \mathbf{y}) = \frac{1}{k} \sum_{i=1}^{k} (x_i - y_i)^2. \tag{7.30}$$

The algorithm starts with an initial codebook $C = \{\mathbf{r}_i; i = 1, \ldots, N\}$ containing N representative vectors $\mathbf{r}_i$ and a training set $T = \{\mathbf{x}_j; j = 1, \ldots, M\}$ of size $M \gg N$. In the first step for T a minimum distortion partitioning P is computed as

$$\begin{aligned} P &= \{S_i; i = 1, \ldots, N\} \\ S_i &= \{\mathbf{x} \mid d(\mathbf{x}, \mathbf{r}_i) \leq d(\mathbf{x}, \mathbf{r}_j), \quad \forall j \neq i \ \wedge \ \mathbf{x} \notin S_k \ \forall k < i\}. \end{aligned} \tag{7.31}$$

Partition S_i contains all vectors $\mathbf{x}$ which can be reproduced by $\mathbf{r}_i$ with least possible distortion. Ties are broken by favoring the partition with the smallest index. P is then the best partition for codebook C, but unfortunately C is not optimal. To obtain the best codebook C^* for our partitioning P we compute

for each partition S_i a centroid $\mathbf{r}_i^*$ which serves as its new representative vector:

$$C^* = \{\mathbf{r}_i^* \mid \mathbf{r}_i^* = \frac{1}{|S_i|} \sum_{j:\mathbf{x}_j \in S_i} \mathbf{x}_j\}, \tag{7.32}$$

where $|S_i|$ denotes the number of vectors belonging to partition S_i. Since we have obtained a new codebook C^* we need to recompute the partitioning of our training set. Therefore we set C to C^* and repeat the partitioning (7.31) and centroid calculation (7.32) until the average distortion

$$D_{\text{avg}} = \frac{1}{M} \sum_{j=1}^{M} \min_{\mathbf{r} \in C} d(\mathbf{x}_j, \mathbf{r}) \tag{7.33}$$

obtained with codebook C cannot be significantly decreased by further iterations.

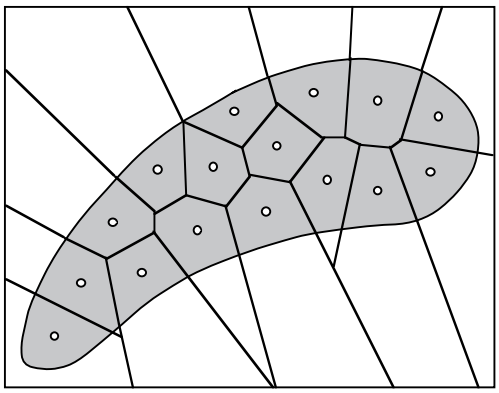

Figure 16 2-D vector space with uniformly distributed data quantized by full search LBG VQ.

Fig. 16 shows how a two-dimensional vector space is partitioned by applying the LBG algorithm for $N = 16$ using (7.30) as the distortion measure. The training set is represented by the shaded banana-shaped region wherein training vectors are uniformly distributed. White dots stand for the representatives of the final codebook.

A disadvantage of the codebook found by the LBG algorithm is, that it is unstructured. This implies that a codevector can only be found by a full search algorithm which requires roughly kN subtractions, multiplications, and additions. On the other hand, decoding is simply a table look-up. We will compare

unstructured vector quantizers for fixed word length encoding using the LBG algorithm and the mean squared error distortion measure for subband image coding in subsection 3.3.

Entropy-Constrained VQ

The LBG algorithm can be extended for codebook design under an entropy constraint [53]. Entropy-constrained vector quantization (ECVQ) minimizes distortion for a given average codeword length rather than a given codebook size. The average codeword length constraint is built into the codebook design by a Lagrange multiplier formulation. Instead of just minimizing the average distortion D, now $D + \lambda R$ is minimized, where R is the average bit-rate.

Let $l(i)$ denote the codeword length needed to encode the representative vector $\mathbf{r}_i$. Then the Lagrange formulation can be easily incorporated into the LBG algorithm described in the previous section. The LBG rule (7.31) for computing the minimum distortion partition P now contains an entropy constraint:

$$\begin{aligned} P &= \{S_i;\ i = 1, \ldots, N\} \qquad (7.34) \\ S_i &= \{\mathbf{x} \mid d(\mathbf{x}, \mathbf{r}_i) + \lambda l(i) \leq d(\mathbf{x}, \mathbf{r}_j) + \lambda l(j), \forall j \neq i \ \wedge \ \mathbf{x} \notin S_k \ \forall\, k < i\}. \end{aligned}$$

A vector $\mathbf{x}$ will be mapped to partition S_i if the distortion between $\mathbf{x}$ and $\mathbf{r}_i$ biased by the codeword length $l(i)$ is minimized. Ties are broken as in (7.31). In other words, the best representative for $\mathbf{x}$ now is the one that yields the best tradeoff between distortion and codeword length. Computation of the optimal codebook C^* for P is carried out according to (7.32). The original LBG algorithm for fixed codeword length is a special case with $\lambda = 0$.

In [53] it is proposed to use $l(i) = \log_2(\frac{1}{p(i)})$ as the "ideal" codeword length for $\mathbf{r}_i$ with $p(i) = \frac{M}{|S_i|}$. This assumes that noninteger codeword lengths are allowed. Another possibility incorporates the construction of a Huffman code [54] into the codebook design. In [53] it is reported that assuming noninteger codeword lengths during codebook design performs nearly identical to a system with an incorporated Huffman algorithm.

Fig. 17 shows the changed subdivision of the vector space compared to Fig. 16 if an entropy-constraint is built into the LBG algorithm. Since the codeword length is taken into account in the codebook design, the decision hyperplanes separating the regions in vector space with the smallest Euclidean distance around individual representatives are shifted compared to the fixed word length VQ. As an example, consider an input vector that is approximately equally close

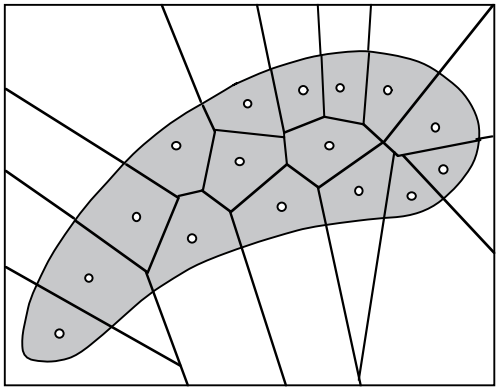

Figure 17 2-D vector space with uniformly distributed data quantized by full search LBG VQ with entropy-constraint.

to two reproduction vectors. It is advantageous to choose the representative with the shorter codeword length, even though its associated distortion might be slightly larger. Since there are more bits remaining, the resulting overall distortion will be smaller. When designing an ECVQ codebook, the initial codebook size has to be sufficiently large. Otherwise, the resulting codebook will not achieve the theoretical minimum distortion for a given rate [53]. Redundant vectors are automatically eliminated in the iterative design process.

Subband image coding with ECVQ codebooks is compared to other quantization schemes in subsection 3.3.

Lattice VQ

Another important class of vector quantizers is lattice quantizers, which are characterized by a highly regular codebook structure. We say that a regular arrangement of points in n-dimensional space is a lattice, Λ_n. A quantizer that uses the points of a lattice as codevectors is said to be a lattice vector quantizer. Compared to VQ with unstructured codebooks, lattice VQ has a much lower encoder complexity at interesting bit-rates. Fast quantization and decoding algorithms for several lattices can be found in [55] and [56]. As an example we describe later in this section a fast quantization algorithm for an eight-dimensional lattice. In subsection 3.3, we shall present results for lattice vector quantizers based on various lattices.

An n-dimensional lattice, Λ_n, is defined as

$$\Lambda_n = \{\mathbf{x} \in R^m \mid \mathbf{x} = \mathbf{zG}, \mathbf{z} \in Z \times Z \times \ldots \times Z = \mathbf{Z}^n\} \tag{7.35}$$

where $\mathbf{G}$, the generator matrix of the lattice, is an $m \times n$ matrix ($m \geq n$) with n linear independent rows belonging to R^m. Some lattices can be defined in a more convenient way. For example, for $n \geq 3$ the D_n lattice can be described as

$$D_n = \{\mathbf{x} \in Z^n \mid \sum_{i=1}^{n} x_i \ \text{ even}\}. \tag{7.36}$$

It contains all points of the Z^n-lattice whose coordinate sum is even. Lattices can often be expressed as a union of cosets of simpler lattices. A coset of a lattice is a lattice resulting from adding a fixed vector to each lattice point. For instance the E_8-lattice can be defined as

$$E_8 = D_8 \ \bigcup \ \left[D_8 + (\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2})\right]. \tag{7.37}$$

which is the union of the D_8-lattice and a coset of the D_8-lattice obtained by adding $\mathbf{\frac{1}{2}} = (\frac{1}{2}, \ldots, \frac{1}{2})$ to each lattice point. The E_8-lattice may also be defined by the generator matrix

$$\mathbf{G}_{E_8} = \begin{bmatrix} 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \end{bmatrix}. \tag{7.38}$$

Another important lattice is the A_n-lattice. It can be defined by the $(n+1) \times n$–dimensional generator matrix

$$\mathbf{G}_{A_n} = \begin{bmatrix} -1 & 1 & 0 & 0 & \ldots & 0 & 0 \\ 0 & -1 & 1 & 0 & \ldots & 0 & 0 \\ 0 & 0 & -1 & 1 & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \ldots & -1 & 1 \end{bmatrix}. \tag{7.39}$$

It can be shown, that for a smooth multidimensional probability density function and sufficiently fine quantization, a lattice vector quantizer derived from the densest multidimensional sphere packing can approach the minimum mean squared quantization error for a given entropy [57]. Therefore lattices of interest are mainly those corresponding to densest sphere packings. For instance, in two dimensions, it is known that a hexagonal partitioning of the space corresponds to the densest packing. The corresponding hexagonal lattice is called the A_2-lattice (see Fig. 18). In 3-dimensional space, the A_3-lattice is known to yield the densest sphere packing. Unfortunately in higher dimensions, lattices corresponding to densest sphere packings are currently only known for 8 dimensions (E_8-lattice) and 24 dimensions (Leech lattice).

Let us now consider an example of a fast lattice quantization algorithm that finds the D_n-lattice point closest to a given input vector. For $x \in R$, let $f(x)$ be defined as the closest integer to x and $w(x) = f(x) + \text{sign}(x - f(x))$ with $\text{sign}(y) = 1$ if $y \geq 0$ and $\text{sign}(y) = -1$ otherwise. The D_n-lattice point closest to $\mathbf{x} \in R^n$ can be found in the following way: compute first $f(\mathbf{x}) = (f(x_1), \ldots, f(x_n))$, the point of Z^n which is closest to $\mathbf{x}$. If $f(\mathbf{x})$ does not have an even sum of components compute $g(\mathbf{x}) = (f(x_1), \ldots, w(x_i), \ldots, f(x_n))$, where x_i is the component furthest away from an integer. Whichever of $f(\mathbf{x})$ and $g(\mathbf{x})$ has an even sum of coordinates is the closest D_n-lattice point to $\mathbf{x}$.

If $\Phi_\Lambda(\mathbf{x})$ is the closest lattice point for a given lattice Λ then $\Phi_\Lambda(\mathbf{x} - \mathbf{t}) + \mathbf{t}$ is the closest point on the translated lattice $\Lambda + \mathbf{t}$. With this insight we now can formulate algorithms for lattices that can be defined as the union of lattice cosets, e.g. for the E_8-lattice. Knowing the algorithm for the D_8-lattice we can find a nearest E_8-lattice point by computing $\Phi_{E_8}(\mathbf{x})$ and $\Phi_{E_8}(\mathbf{x} - \frac{\mathbf{1}}{\mathbf{2}}) + \frac{\mathbf{1}}{\mathbf{2}}$ and then choosing among the two results the one closest to $\mathbf{x}$. This can be generalized for every lattice that can be described as a union of cosets.

The structure of a lattice gives an implicit partitioning of the vector space into lattice cells that are regular in contrast to the partitionings found by the LBG algorithm. We have already seen that the best codebook for a set of given partitions results from using each partition's centroid as its representative (see 7.32). This also can be applied to lattice quantizers and motivates the use of centroids as codevectors for lattice cells. The centroids are computed for a training set together with lattice cell probabilities and are stored in a special codebook. The probabilities can be used to construct for instance a Huffman code for the encoding of lattice cells. To keep memory requirements in reasonable bounds, centroids and probabilities should only be stored for the most popular lattice cells. In the other lattice cells the lattice point itself is used for reconstruction and equal probabilities are assumed.

This is illustrated in Fig. 18 for an A_2 lattice vector quantizer. The use of the centroid instead of a lattice point is best seen for cells that are only partially covered by the shaded training set.

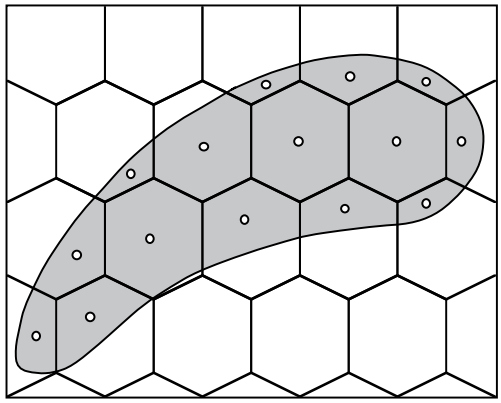

Figure 18 2-D vector space with uniformly distributed data quantized by A_2-lattice VQ.

3.3 Experimental Comparison of Vector Quantizers for Image Subbands

In this section, various vector quantization techniques as described in the previous section are compared experimentally, using both synthetic source data and natural images decomposed by a subband pyramid. In each case, we have measured the distortion over a range of bit-rates. Comparisons for other quantization techniques like pyramid vector quantization can be found in [4].

Fig. 19 shows the performance of 8-dimensional VQ for a first-order Gauss-Markov source with a correlation coefficient of $r = 0.95$. This corresponds to the typical correlation of adjacent samples in a baseband image. For comparison, the Shannon lower bound of the rate-distortion function (7.13) is shown in Fig. 19 as well. Unstructured VQ with or without variable length coding both perform well. It is somewhat surprising that ECVQ does not achieve any gain over fixed codeword length (CWL) VQ. The E_8 lattice vector quantizer is about 2 dB worse than unstructured VQ. Optimum scalar quantization followed by entropy coding is also shown. It does not exploit the correlation between successive samples and thus performs poorly.

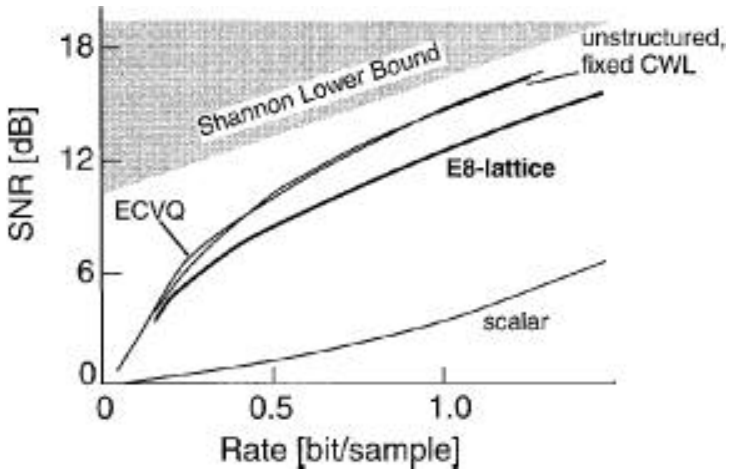

Figure 19 Distortion rate curves for 8-D vector quantization of a Gauss-Markov source with correlation coefficient $r = 0.95$.

Subband images are much less correlated than the full band image. Their amplitude distribution resembles a Laplacian distribution. We have therefore compared the various VQs for data from a memoryless Laplacian source. Results for 8-D VQ are shown in Fig. 20. The curves are dramatically different from the results shown in Fig. 19. Since the source is memoryless, gains beyond scalar quantization are small. Unstructured fixed word length VQ performs poorly at low bit-rates. E_8-lattice VQ performs as well as unstructured ECVQ at low bit-rates.

The following results have been obtained for the well-known image "Lenna" (of size 512×512), decomposed in a subband pyramid. The vector quantizers were obtained using a training set of four images. "Lenna" was not contained in the training set. We have measured distortion rate curves for each individual subband for each of the vector quantizers. It is beyond the scope of this chapter to present these results comprehensively. Rather, we show two typical examples in Figs. 21 and 22. The presentation has been changed compared to Figs. 19 and 20. It is more convenient to work with mean squared error curves rather than SNR for optimum bit allocation, as discussed in the next section.

Fig. 21 shows results for 8-D vector quantization of the horizontal high band of the first pyramid level. Unstructured VQ with fixed word length performs even worse than optimum scalar quantization with (memoryless) entropy coding. As expected, unstructured VQ with entropy coding performs best. The E_8-lattice

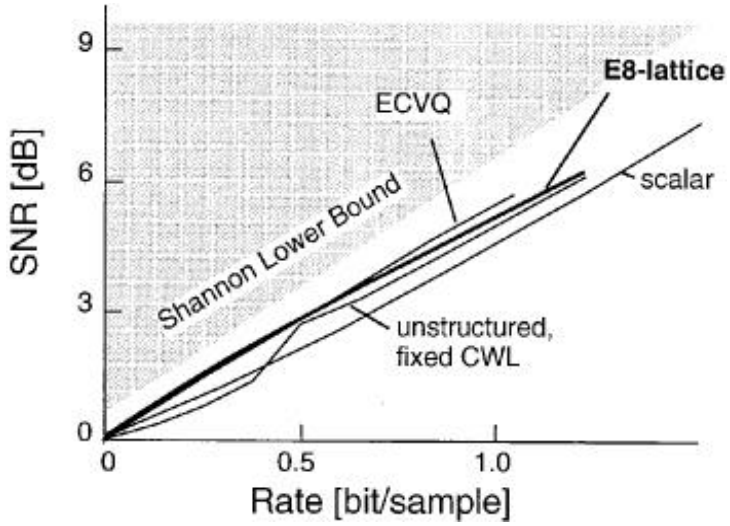

Figure 20 Distortion rate curves for 8-D vector quantization of a memoryless Laplacian source.

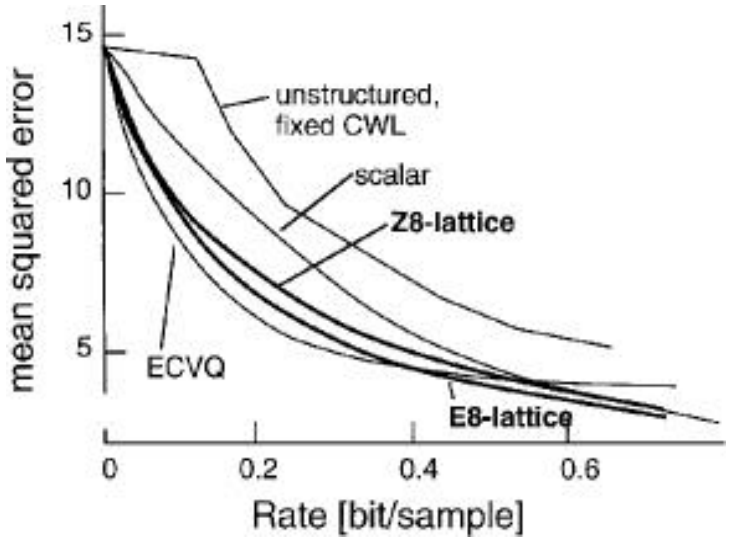

Figure 21 Distortion rate curves for 8-D vector quantization of the horizontal high frequency band of the first pyramid level.

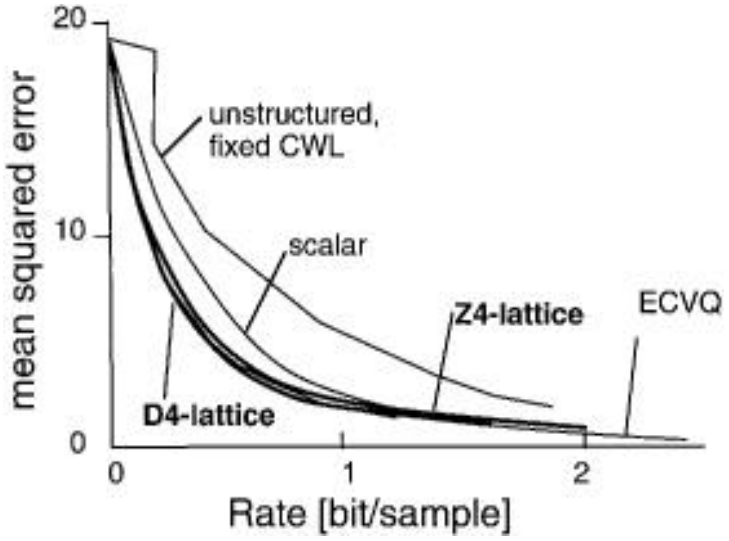

Figure 22 Distortion rate curves for 4-D vector quantization of the vertical high frequency band of the second pyramid level.

VQ is close to the optimum. At higher rates the training set becomes too small to reliably train all the degrees of freedom of the unstructured ECVQ codebook, and its performance becomes slightly worse than E_8-lattice VQ. Interestingly, the simple orthogonal Z^8-lattice performs almost as well as the E_8-lattice. Since it requires even less computation than E_8, it is certainly a very interesting alternative. Fig. 22 shows similar results for 4-D VQ of the vertical high band of the second pyramid level. The unstructured ECVQ performs best, but there is practically no loss when a highly structured D_4-lattice VQ or an even simpler orthogonal Z^4-lattice VQ is used. VQ with fixed word length does not perform well at all, and it is even worse than optimum scalar quantization. A more detailed discussion of these results has been presented in [40].

3.4 Bit-Rate Allocation

A subband image coder can be regarded as a special case of the system shown in Fig. 23. An original image x is represented by M independent bit-streams with individual bit-rates R_i, yielding an overall rate

$$R = R_0 + R_2 + ... + R_{M-1}. \tag{7.40}$$

At the decoder, a signal $\hat{x}$ is reconstructed that deviates from x by some average distortion D.

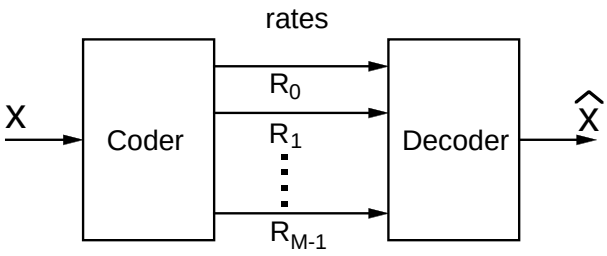

Figure 23 Codec representing signal x by M parallel independent bit-streams.

We assume that the multivariate distortion rate function $D(R_0, R_1, ..., R_{M-1})$ is strictly convex and differentiable everywhere, and that

$$\frac{\partial D}{\partial R_i} \leq 0, \forall i \tag{7.41}$$

i.e., increasing the rate for any one of the bit-streams will decrease the distortion D. We can find the optimum bit allocation that minimizes D subject to a fixed overall rate R by setting the total differential of the distortion to zero, i.e.

$$dD = \frac{\partial D}{\partial R_0} dR_0 + \frac{\partial D}{\partial R_1} dR_1 + ... + \frac{\partial D}{\partial R_{M-1}} dR_{M-1} = 0. \tag{7.42}$$

From (7.40), we obtain

$$dR_0 + dR_1 + \ldots + dR_{M-1} = 0. \tag{7.43}$$

From (7.42) and (7.43) follows the optimum bit allocation condition

$$\frac{\partial D}{\partial R_0} = \frac{\partial D}{\partial R_1} = \ldots = \frac{\partial D}{\partial R_{M-1}}. \tag{7.44}$$

Eq. (7.44) can be interpreted as follows. If we have one more (infinitesimal) bit to spend, we would add it to the bit-stream with the smallest $\frac{\partial D}{\partial R_i}$, since this would decrease distortion by the greatest amount. We would continue to do so, until some other bit-stream offers a greater pay-off. The optimum balance is achieved according to (7.44), when it does not matter to which bit-stream we add the next bit.

If the overall distortion D is the sum of the individual subband distortions D_i, i.e.,

$$D = \sum_i D_i, \tag{7.45}$$

(7.44) can be simplified to

$$\frac{\partial D_i}{\partial R_i} = \frac{\partial D_j}{\partial R_j} \qquad \text{for all } i, j. \tag{7.46}$$

In other words, we have to pick points of equal slope on the individual distortion rate curves for each subband. Equation (7.45) holds if mean-squared error is used as the distortion measure and the subbands are orthogonal. Many subband decomposition schemes are designed with an orthogonality criterion or are nearly orthogonal. Based on (7.46), we have optimized a subband pyramid

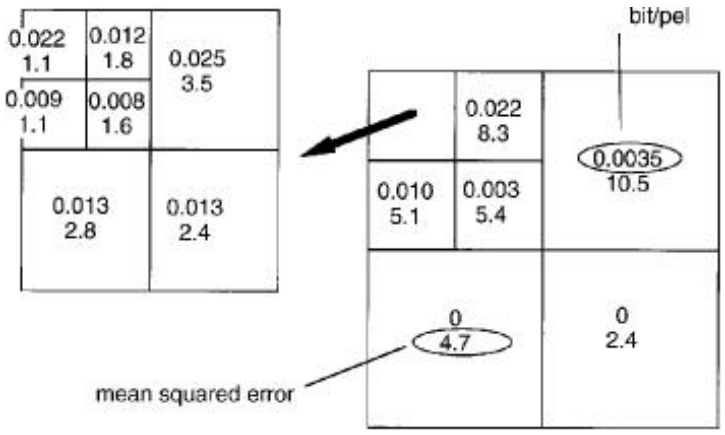

Figure 24 Bit allocation and mean squared error contributions of individual subbands.

coder that uses E_8-lattice VQ on the first pyramid level, D_4-lattice VQ on the second level, and the A_2-lattice on the third and fourth levels. With this coder we have compressed test image "Lenna" and have achieved a bitrate of 0.136 bpp while maintaining a peak-to-peak SNR of 30.9 dB. For a visual impression of the achieved quality, refer to [40]. The overall quality is very good for a compression ratio of approximately 60:1 (starting at 8 bpp). Fig. 24 summarizes the bit allocation among subbands and the contribution of each subband to the overall distortion. Interestingly, the mean squared error (MSE) contribution of the various subbands differs considerably. It ranges from MSE = 1.1 for the lowest band to MSE = 10.5 for the horizontal high band on

the first pyramid level. If we do not use rule (7.46) for bit allocation but rather allocate bits for equal mean squared error contribution of the bands (a popular text book solution motivated by the rate-distortion analysis presented in section 1.1), overall SNR drops by 0.8 dB for the same bit-rate. We also encoded the image with orthogonal lattice quantizers, and it was confirmed that these quantizers perform only slightly worse than the denser lattices. The excellent performance of subband pyramid coding in combination with lattice VQ has been confirmed by results reported by Barlaud et. al. [58] [59].

4 CONCLUSIONS

In this chapter, we have reviewed the principles of subband coding of still images and discussed a variety of algorithms employed in practical systems today. Compression can be achieved by exploiting the properties of both signal source and signal receiver. Rate-distortion theory establishes the relationship between the fidelity of the reconstructed image and the lowest transmission bit-rate for a given source model and a given distortion measure. An efficient coder jointly encodes many symbols; ideally, the reconstruction error is statistically independent from the reconstructed signal. For an ergodic Gaussian source with memory it is optimal to split the signal into frequency bands of (ideally) infinitesimal bandwith and to encode those spectral components independently, according to the signal energy contained in the band. This is the underlying idea and the motivation for subband image coding.

Subband image coders consist of three essential stages: decomposition of the signal into frequency bands by means of subband filter banks, quantization of the subband signals, and (lossless) entropy coding.

For perfect reconstruction, subband filter banks have to obey certain design rules. Quadrature mirror filters (QMF) with linear phase can yield nearly perfect reconstruction and are widely used in image coding. The signal may be split into bands of equal or unequal bandwiths by appropriate cascading of two–band decompositions. Subband pyramids that recursively decompose the low band are now widely used. Another important decomposition is the discrete cosine transform (DCT), which is in fact a special case of a subband decomposition.

After the decomposition stage, the subband signals have to be compressed. A scalar quantizer quantizes each sample independently and is easy to implement.

Better results are achieved when jointly quantizing several samples by means of vector quantization (VQ). Fixed-wordlength vector quantizers can be designed based on a training set using the LBG algorithm. The LBG algorithm can be extended to the design of vector quantizers with an entropy constraint. For a given entropy and a smooth probability density function, a lattice VQ can approach the minimum mean squared error. In subsection 3.3, we have compared various vector quantizers experimentally. For the lowpass band (baseband image), VQ is far superior to scalar quantization. Unstructured VQ with variable length coding is the best choice for encoding of the subband signals. The subband signals are, however, spatially much less correlated, and VQ does not offer as much gain as for the lowpass band. For low bit-rates, unstructured VQ with an entropy constraint (ECVQ) performs best, but it is also computationally most expensive. Lattice VQ is computationally less expensive and performs almost as well.

Finally, we have reviewed conditions for optimum bit allocation among the subbands. In a coding example, we have demonstrated that a monochrome still image can be compressed by $60:1$ while maintaining an acceptable level of image quality.

REFERENCES

[1] CCIR. Encoding Parameters of Digital Television for Studios, CCIR Recommendation 601, 1982.

[2] A.N. Akansu and R.A. Haddad. *Multiresolution signal decomposition.* Academic Press, San Diego, 1992.

[3] J.W. Woods and S.D. O'Neil. Subband coding of images. *IEEE Trans. Acoust. Speech Signal Process.*, ASSP-34:1278–1288, October 1986.

[4] M.E. Blain and T.R. Fischer. A comparison of vector quantization techniques in transform and subband coding of imagery. *Signal Processing: Image Communication*, 3(1):91–105, February 1991.

[5] N. Ahmed, T. Natarajan, and K.R. Rao. Discrete cosine transform. *IEEE Transactions on Computers*, C-23:90–93, January 1974.

[6] ISO 10918-1 JPEG Draft International Standard, CCITT Recommendation T.81, 1992.

[7] ISO/IEC 13818-2, Generic Coding of Moving Pictures and Associated Audio, Recommendation H.262 (MPEG-2), March 1994. Draft International Standard.

[8] CCITT. Video codec for audiovisual services at p x 64 kbit/s, CCITT recommendation H.261, 1990.

[9] P.P. Vaidyanathan. *Multirate Systems and Filter Banks.* Prentice Hall, Englewood Cliffs, New Jersey, 1993.

[10] U. Horn and B. Girod. Pyramid coding using lattice vector quantization for scalable video applications. In *Proc. PCS '94*, Sacramento, September 1994.

[11] B. Girod. Scalable video for multimedia workstations. *Computers + Graphics*, 17(3):269–276, 1993.

[12] D. Taubman and A. Zakhor. Rate and resolution scalable subband coding of video. In *Proceedings ICASSP 94*, pages V–493 – V–496, 1994.

[13] M. Vetterli and K.M. Uz. Multiresolution coding techniques for digital television: A review. *Multidimensional Systems and Signal Processing*, 3:161–187, 1992.

[14] Robert G. Gallagher. *Information Theory and Reliable Communication.* John Wiley and Sons, 1968.

[15] T. Berger. *Rate Distortion Theory.* Prentice-Hall, 1971.

[16] R.E. Blahut. Computation of channel capacity and rate distortion functions. *IEEE Transactions on Information Theory*, IT-18:460–473, 1972.

[17] S. Nanda and W.A. Pearlman. Tree coding of image subbands. *IEEE Trans. on Image Proc.*, 1(2):133–147, April 1992.

[18] M.J.T. Smith and S.L. Eddins. Analysis/synthesis techniques for subband image coding. *IEEE Trans. Acoust., Speech, Signal Processing*, ASSP-38:1446–1456, August 1990.

[19] J.D. Johnston. A filter family designed for use in quadrature mirror filter banks. In *Proceedings ICASSP 80*, pages 291–294, 1980.

[20] M. Vetterli. Multi-dimensional sub-band coding: some theory and algorithms. *Signal Processing*, 6:97–112, April 1984.

[21] H. Gharavi and A. Tabatabai. Subband coding of digital image using two-dimensional quadrature mirror filtering. In *Proceedings SPIE Visual Communications and Image Processing*, pages 51–61, 1986.

[22] T. Kronander. A new approach to recursive mirror filters with a special application in subband coding of images. *IEEE Trans. Acoust. Speech Signal Process.*, ASSP-36(9), September 1988.

[23] D. LeGall and A. Tabatabai. Sub-band coding of digital images using symmetric short-kernel filters and arithmetic coding techniques. In *Proc. ICASSP '88*, pages 761–764, 1988.

[24] Edward H. Adelson and Eero P. Simoncelli. Subband image coding with three-tap pyramids. In *Proceedings Picture Coding Symposium '90*, 1990.

[25] J.H. Husøy. Low-complexity subband coding of still images and video. *Optical Engineering*, 30(7), July 1991.

[26] H.S. Malvar. *Signal Processing with Lapped Transforms.* Artech House, 1992.

[27] Eero P. Simoncelli and Edward H. Adelson. Non-separable extensions of quadrature mirror filters to multiple dimensions. *Proceedings of the IEEE*, 78(4):652–664, April 1990.

[28] R.H. Bamberger and M.J.T. Smith. A filter bank for the directional decomposition of images: Theory and design. *IEEE Trans. Signal Process.*, 40(4):882–893, April 1992.

[29] K.R. Rao and P. Yip. *Discrete Cosine Transform.* Academic Press, San Diego, 1990.

[30] J. Pearl. On coding and filtering stationary signals by discrete Fourier transforms. *IEEE Trans. on Information Theory*, pages 229–232, March 1973.

[31] C. Diab, R. Prost, and R. Goutte. Exact image subband decomposition/reconstruction by DCT. *Signal Processing: Image Communication*, 4(6):489–496, November 1992.

[32] F. Bellifemine, C. Cafforio, A. Chimienti, and R. Picco. Combining DCT and subband coding into an intraframe coder. *Signal Processing: Image Communication*, 5(3):235–248, May 1993.

[33] E.A. da Silva and M. Ghanbari. A DCT-based aliasing cancellation method in subband coding. *IEEE Trans. on Circuits and Systems for Video Techn.*, 3(5):384–387, October 1993.

[34] P.J. Burt and E.H. Adelson. The Laplacian pyramid as a compact image code. *IEEE Trans. Commun.*, COM-31:532–540, April 1983.

[35] J.H. Husøy and T.A. Ramstad. Subband coding of images employing an efficient parallel filter bank. In *Proc. SPIE 1199, Visual Communications and Image Processing IV*, pages 752–763, 1989.

[36] L. Vandendorpe. Optimized quantization for image subband coding. *Signal Processing: Image Communication*, 4(1):65–79, November 1991.

[37] J. Lee and B.W. Dickinson. Subband video coding with temporally adaptive motion interpolation. In *Proceedings ICASSP 94*, pages V–497 – V–500, 1994.

[38] P.H. Westerink, D.E. Boekee, J. Biemond, and J.W. Woods. Subband coding of images using vector quantization. *IEEE Trans. on Communication*, 36(6):713–719, June 1988.

[39] Y.H. Kim and J.W. Modestino. Adaptive entropy coded subband coding of images. *IEEE Trans. on Image Proc.*, 1(1):31–48, January 1992.

[40] T. Senoo and B. Girod. Vector quantization for entropy coding of image subbands. *IEEE Trans. on Image Processing*, 1(4):526–532, October 1992.

[41] I. Furukuwa, M. Nomura, and S. Ono. Hierarchical sub-band coding of super high definition image with adaptive block-size multistage VQ. *Signal Processing: Image Communication*, 5(5-6):527–538, December 1993.

[42] R.E. Dyck van and S.A. Rajala. Subband/VQ coding of color images with perceptually optimal bit allocation. *IEEE Trans. on Circuits and Systems for Video Techn.*, 4(1):68–82, February 1994.

[43] N. Mohsenian and N.M. Nasrabadi. Edge-based subband VQ techniques for images and video. *IEEE Trans. on Circuits and Systems for Video Techn.*, 4(1):53–67, February 1994.

[44] M. Antonini, M. Barlaud, P. Mathieu, and I. Daubechies. Image coding using wavelet transform. *IEEE Trans. on Image Processing*, 1:205–220, April 1992.

[45] S.P. Lloyd. Least squares quantization in PCM. *Institute of Mathematical Statistics Meeting, Atlantic City, N.J.*, September 1957.

[46] J. Max. Quantizing for minimum distortion. *IRE Trans. on Information Theory*, pages 7–12, March 1960.

[47] P.F. Panter and W. Dite. Quantization distortion in pulse count modulation with nonuniform spacing of levels. *Proc. IRE*, pages 44–48, January 1951.

[48] R. Wood. On optimum quantization. *IEEE Trans. on Information Theory*, pages 248–252, March 1969.

[49] H. Brusewitz. *Quantization and Entropy Coding.* PhD thesis, The Royal Institute of Technology, Stockholm, Sweden, March 1987.

[50] N.M. Nasrabadi and R.A. King. Image coding using vector quantization: A review. *IEEE Trans. on Communications*, 36(8):957–971, August 1989.

[51] A. Gersho and R.M. Gray. *Vector Quantization and Signal Compression.* Kluwer Academic Publishers, Dordrecht, 1992.

[52] Y. Linde, A. Buzo, and R.M. Gray. An algorithm for vector quantizer design. *IEEE Trans. on Communications*, COM-28(1):84–95, January 1980.

[53] P.A. Chou, T. Lookabaugh, and R.M. Gray. Entropy-constrained vector quantization. *IEEE Trans. on Acoustics, Speech, and Signal Processing*, 37(1):31–42, January 1989.

[54] D.A. Huffman. A method for the construction of minimum-redundancy codes. In *Proc. IRE*, volume 40, pages 1098–1101, September 1952.

[55] J.H. Conway and N.J.A. Sloane. A fast encoding method for lattice codes and quantizers. *IEEE Trans. on Information Theory*, IT-29(6):820–824, November 1983.

[56] J.H. Conway and N.J.A. Sloane. Fast quantizing and decoding algorithms for lattice quantizers and codes. *IEEE Trans. on Information Theory*, IT-28(2):227–232, March 1982.

[57] Y. Yamada, S. Tazaki, and R.M. Gray. Asymptotic performance of block quantizers with difference distortion measures. *IEEE Trans. on Information Theory*, IT-26(1):6–14, January 1980.

[58] M. Barlaud, P. Solé, M. Antonini, and P. Mathieu. A pyramidal scheme for lattice vector quantization of wavelet coefficients applied to image coding. In *Proc. ICASSP '92*, San Francisco, CA, March 1992.

[59] M. Barlaud, P. Solé, T. Gaidon, M. Antonini, and P. Mathieu. Pyramidal lattice vector quantization for multiscale image coding. *IEEE Transactions on Image Processing*, 3(4):367–381, July 1994.